Effect of Eye Testing Order on Automated Perimetry Results Using the Swedish Interactive Threshold Algorithm Standard 24-2

Yaniv Barkana, MD; Yariv Gerber, PhD; Ricardo Mora, MD; Jeffrey M. Liebmann, MD; Robert Ritch, MD

Objective: To evaluate whether the order of eye testing affects the mean deviation (MD) or the test reliability of visual field testing using the Swedish Interactive Threshold Algorithm (SITA) standard 24-2.

Methods: Consecutive patients with manifest or suspect glaucoma with 2 prior sets of SITA standard 24-2 test results performed on the right eye first were enrolled. A subsequent test was performed on the left eye first. For each eye, the MD and the test reliability indexes (≥20%) were compared among the 3 successive examinations.

Results: Forty-seven patients (29 women and 18 men; mean±SD age, 70.6±11.9 years) were enrolled. The MD±SD was −5.83±5.43 dB OD and −5.46±4.86 dB OS. There was no statistically significant difference in the MD or the test reliability among the 3 test results for either eye. Fixation loss was responsible for the unreliable fields in almost all cases.

Conclusions: Among this cohort of patients experienced with automated perimetry in a glaucoma subspecialty practice, changing the order of eye testing using the SITA standard 24-2 did not have a significant effect on the MD or the test reliability. Intereye fatigue may not be clinically significant with this algorithm. Fixation loss remains a problem with the use of this algorithm.

Arch Ophthalmol. 2006;124:781-784

ACCURATE AUTOMATED VISUAL field (VF) testing is critical for the diagnosis and management of patients with glaucoma. Because it is a subjective psychophysical test, an interpretation of the results must consider patient performance. Patient reliability may be determined by automated assessment of the rate of fixation loss, false-positive and false-negative responses. Fixation may also be assessed through direct observation by the perimetrist. The reliability indexes are frequently outside the recommended fixation loss and false-positive and false-negative limits, complicating the interpretation of test results.1,2 True progression must be differentiated from artifactitious change caused by variable patient performance and other factors.3

Fatigue during the course of the VF examination may adversely affect patient performance. Light sensitivity has been shown to decrease progressively during automated perimetry examination of a single eye or in comparing the results of the second eye with those of the first eye.3,4 When designing studies that involve automated perimetry, the fatigue effect is compensated for by allowing sufficient rest between tests or by randomizing the order of eye testing.6-8

It is common practice to begin testing with the right eye. If the examiner is more interested in the results of the left eye, he or she may begin testing with the left eye. However, to our knowledge, the effect of this practice on test results is not established. Previous studies9,10 of fatigue during a single examination used older longer-duration testing algorithms. The use of the Swedish Interactive Threshold Algorithm (SITA) standard 30-2 requires about half the time of the older standard 30-2 algorithm. The commonly used SITA standard 24-2 is even shorter in duration and may be associated with less fatigue. The objective of this study was to determine whether the mean deviation (MD) or the test reliability changed for either eye based on whether it was the first eye or the second eye to be examined.

METHODS

Consecutive patients with manifest or suspect glaucoma seen by one of us (J.M.L.) in a glaucoma subspecialty practice during a
RESULTS

Forty-seven patients (29 women and 18 men; mean ± SD age, 70.6±11.9 years) were enrolled. Among right eyes, 28 had manifest glaucoma and 19 had suspect glaucoma. Among left eyes, 30 had manifest glaucoma and 17 had suspect glaucoma.

The mean ± SD visual acuity during the 3 tests was 6.39±1.29, 6.32±1.16, and 6.17±0.99 minutes OD and 6.34±1.10, 6.42±1.00, and 6.34±1.27 minutes OS. There was no statistically significant difference in the test duration among the 3 test results for either eye.

The MD±SD among the 3 test results was −5.83±5.43 dB OD and −5.46±4.86 dB OS. There was no statistically significant difference in the MD among the 3 test results for either eye (Table).

Unreliable fields (with ≥1 reliability index ≥20%) were obtained in 29.8% of right eyes and in 32.6% of left eyes. Fixation loss was responsible for the unreliable fields in almost all cases (85.7% of right eyes and 95.7% of left eyes). There was no statistically significant difference in the percentage of test results with fixation loss of at least 20% among the 3 test results for either eye (Table).

To further characterize the significance of the order of eye testing, we tested whether individual patients consistently performed reliably during testing of the first eye and consistently performed unreliably during testing of the second eye. Based on fixation loss rates during right eye VF tests, 3 of 47 subjects performed reliably on both tests when the right eye was tested first and performed unreliably when the right eye was tested second. Based on fixation loss rates during left eye VF tests, there was no case in which a reliable test result was obtained when the left eye was tested first and unreliable test results were obtained on both tests when the left eye was tested second.

The Table gives the percentages of test results that were deemed unreliable because of at least 20% false-positive or false-negative responses. There were consistently few such test results.

COMMENT

Fatigue during a single test of automated perimetry is believed to affect test results. Clinicians often instruct their technicians to test the eye of interest first, believing that patient performance is poorer during testing of the second eye. However, there is a scarcity of literature quantifying this effect, especially among patients with glaucoma; furthermore, the existing literature evaluated older

Table. Results of the 3 Visual Field Tests*

<table>
<thead>
<tr>
<th>Eye</th>
<th>Test 1</th>
<th>Test 2</th>
<th>Test 3</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OD</td>
<td>−5.45±5.11</td>
<td>−5.72±5.38</td>
<td>−6.31±5.81</td>
<td>.66</td>
</tr>
<tr>
<td>OS</td>
<td>−5.41±5.11</td>
<td>−5.73±4.71</td>
<td>−5.23±4.77</td>
<td>.93</td>
</tr>
</tbody>
</table>

*Data are given as percentages unless otherwise indicated.
longer-duration algorithms of automated perimetry.\textsuperscript{3-5} More recent studies used the full-threshold 30-2 algorithm, which requires approximately 15 minutes per eye, whereas the current SITA standard 24-2 typically requires less than half of this time.\textsuperscript{9,11}

Brenton et al\textsuperscript{12} evaluated 20 healthy subjects using the 30-2 algorithm of the Humphrey field analyzer, allowing rest whenever needed and no more than 8 minutes of testing without a break. The mean interocular sensitivity difference (right eye minus left eye) ranged from –0.5 to 1.1 dB. Among their cohort, the authors concluded that even if there was a slight fatigue effect it was counterbalanced by a learning effect, enhancing the performance during the second eye test.

Two studies simulated tests lasting 15 minutes per eye. In the first study, Hudson et al\textsuperscript{4} used the Octopus perimeter to test 20 healthy subjects and 20 subjects with ocular hypertension. The test duration for each eye was approximately 15 minutes, during which 59 locations within 28.3° were tested (phase 1) and then retested (phase 2), so that a comparison of the 2 threshold results would provide a measure of fatigue effect. There was a 1-minute break between phases and a 3-minute break between eyes. The order of eye testing was randomized for each subject. The authors found that for each eye there was a statistically significant poorer MD (ie, reduction in sensitivity) between the 2 phases and there was a statistically significant poorer MD in the second eye vs the first eye. Similar results were found for loss of variance.

In the second study, Searle et al\textsuperscript{5} tested 38 healthy perimetrically naive subjects using the Humphrey field analyzer. Thirty locations were thresholded in each eye during each of 3 successive 5-minute periods at each of 2 visits separated by 2 weeks. This permitted the comparison of the sequential sensitivity at the same locations during a simulated 15-minute test. At both visits, statistically significant decreases in mean sensitivity were observed during the 3 phases for each eye and between the first and second eyes. The differences ranged from 1.1 to 2.0 dB. No trends were observed for the reliability indexes among the 3 phases, between the first and second eyes, or between visits.

The authors of these studies\textsuperscript{4,5} suggested that the normative database for automated perimeters should reflect the order of eye testing and that values should differ between the 2 eyes. The results of our study do not support this suggestion with reference to the SITA standard 24-2. We found no significant MD difference when eyes were tested first or second.

Two other studies examined fatigue effect during testing of a single eye. In the first study, Heijl and Drance\textsuperscript{3} tested 21 patients with glaucoma using the Competer automated perimeter with a program that tested 6 locations repeatedly during 30 minutes. They found a gradual decline in sensitivity, which was more pronounced in locations close to or within the VF defects.

In the second study, Marra and Flammer\textsuperscript{13} used a much shorter program that tested a few locations. Using the Octopus 201 perimeter, they tested 100 subjects, of whom 16 had early glaucoma and 32 had prior experience with automated or manual perimetry. The program tested threshold light sensitivity at 3 locations 12 consecutive times, lasting between 5 and 10 minutes per eye. The authors found that the sensitivity remained stable at the 3 locations throughout the examination for most eyes among perimetrically naive and experienced patients.

The reliability of Humphrey field analyzers often complicates the interpretation of test results. Birt et al\textsuperscript{1} analyzed the reliability indexes of 384 VF test results (in both eyes) among 106 patients with glaucoma in an urban tertiary care practice. The tests included 30-2 and 24-2 programs and averaged 11.7 minutes in duration. Fixation loss of at least 20% led to the classification of 38.5% of all fields as being unreliable. False-positive and false-negative errors greater than 33% were less often the cause of unreliable field classification, with rates of 5% and 9%, respectively. Increased test duration correlated significantly with all 3 catch trial errors. Other significant correlates were age and the MD.

Among 51 healthy subjects, 337 patients with ocular hypertension, and 55 patients with glaucoma, Katz et al\textsuperscript{2} analyzed the reliability of the results of 30-2 tests (in single eyes) repeated 3 or 4 times and separated by a mean of 15 months. Fixation loss of at least 20% accounted for 23.5% of the unreliable fields, while false-positive errors (4.1%) and false-negative errors (6.9%) greater than 33% were less common. A similar proportion of subjects had unreliable test results at each visit, with no trend for any of the 3 patient groups. Patients with glaucoma were more likely to have repeated high rates of false-negative responses, suggesting that false-negative responses are more indicative of glaucoma than of patient reliability. A similar observation was reported by Bengtson and Heijl.\textsuperscript{14}

Shorter test duration, with less associated fatigue, might result in increased test reliability among patients. Even with the shorter SITA standard 24-2 program, our results among a population of perimetrically experienced patients indicate that fixation loss remains a significant problem. Although in this study we found no significant difference in fixation loss associated with the order of eye testing, each patient requires individual management. Unreliable examination results because of high rates of false-positive or false-negative responses were uncommon. This is similar to the findings of Birt et al\textsuperscript{1}.

The source of fatigue during automated perimetry is unclear. A decline in concentration and attention during a monotonous task is plausible. Physiologic phenomena during single-eye examinations such as the Ganzfeld blank out and the Troxler phenomenon have been proposed.\textsuperscript{4,15} Whatever the cause, it is likely that a shorter examination would diminish these effects.

Among perimetrically experienced patients with manifest or suspect glaucoma, our results suggest that changing the order of eye testing with the SITA standard 24-2 should not result in meaningful differences in observed sensitivity and reliability. Although we can postulate that these variables are similarly not affected during the course of a single-eye examination, the present study did not address this issue.

The interval between pairs of VF tests averaged 10 months. Therefore, the results might have been affected by changes in glaucomatous neuropathy. However, based on the findings of recent large studies\textsuperscript{10,17} of patients treated...
for glaucoma, we believe that for our population as a whole the mean VF sensitivity did not change significantly as a result of the glaucomatous process during the study period. The long-term variability in automated perimetry test results could have affected our results because we included in our analysis only 2 results of tests of the right eye first and a single test of the left eye first. This factor could be reduced or eliminated by regression to the mean if a larger number of subjects were tested or if a larger number of VF test results from each participant were analyzed. Finally, our study design had the advantage of VF tests performed under real-life conditions, as opposed to artificial study conditions in which participants perform multiple tests on the same day or within a short study-specific period.

Our study population reflected a glaucoma subspecialty practice and comprised subjects with suspect glaucoma and with a wide severity spectrum of manifest glaucoma. Further study of specific populations, such as patients with advanced glaucoma, is warranted.

The results of our study may be corroborated by comparing the right eye with the left eye perimetry results among a large number of patients with glaucoma whose right eyes had been tested first. Assuming that glaucoma affects both eyes similarly among a large population, we would expect no left-right difference in the MD, pattern standard deviation, or reliability. In our study, the visual acuity in either eye did not differ in a statistically or clinically significant manner among the 3 test results. Therefore, we do not believe that the presence of progressive cataract affected our results in a significant way.

CONCLUSIONS

Among a small cohort of patients with manifest and suspect glaucoma who were experienced in automated perimetry, changing the order of eye testing using the SITA standard 24-2 did not have a significant effect on the MD or the test reliability. Using this algorithm, intereye fatigue may not be clinically significant for most subjects. A few individuals might benefit from changing the order of eye testing, and we recommend in clinical practice that this be considered on an individual patient basis. Fixation loss remains a problem with the use of the short SITA standard 24-2 and was not affected by the order of eye testing in this study.

Submitted for Publication: August 16, 2005; final revision received August 22, 2005; accepted November 29, 2005. Correspondence: Robert Ritch, MD, Department of Ophthalmology, The New York Eye and Ear Infirmary, 310 E 14th St, New York, NY 10003 (ritchmd@earthlink.net).

Author Contributions: Dr Barkana had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Financial Disclosure: None.

Funding/Support: This study was supported in part by the Arthur and Phyllis Bargonetti Research Fund of the New York Glaucoma Research Institute and by the American Physicians Fellowship for Medicine in Israel.

Previous Presentation: This study was presented in part at the ARVO (Association for Research in Vision and Ophthalmology) 2005 Annual Meeting; May 4, 2005; Fort Lauderdale, Fla.

REFERENCES


